Comment on 'Predictions of groundwater PFAS occurrence at drinking water supply depths in the United States' by Andrea K. Tokranov et al., Science 386, 748-755 (2024)

Yuze Hao[1]

In the paper titled 'Predictions of groundwater PFAS occurrence at drinking water supply depths in the United States' by Tokranov et al. (2024), the authors apply an XGBoost machine learning model to predict the occurrence of PFAS (per- and polyfluoroalkyl substances) in groundwater. The model uses a probability threshold of 0.315 to classify PFAS contamination levels, a key parameter for the model's predictive capabilities. While the selection of a threshold is essential in model performance, the authors limit their validation to two threshold values—0.315 and 0.5. This comment seeks to critically examine this choice and its potential limitations.[1]

The authors justify the choice of 0.315 as the threshold based on its optimization for model performance, especially in detecting PFAS presence in groundwater. However, the limited exploration of alternative thresholds raises concerns. While thresholds of 0.315 and 0.5 are discussed in the supplementary materials, there is little justification for why these values were prioritized, nor is there an exploration of intermediate thresholds that might better balance precision and recall. The absence of a more comprehensive threshold analysis leaves a gap in the validation process and may affect the generalizability of the model.

Threshold selection is a critical aspect of machine learning model evaluation, especially when dealing with imbalanced datasets or models aimed at minimizing false negatives or false positives. The lack of broader experimentation with a range of thresholds makes it difficult to assess whether 0.315 represents the optimal cut-off for all potential scenarios.

[1] School of Chemistry and Chemical Engineering & Inner Mongolia Engineering and Technology Research Center for Catalytic Conversion and Utilization of Carbon Resource, Molecules, Inner Mongolia University, 235 West University Street, Hohhot, 010021 China, email: haoyuze2021@outlook.com

More detailed model evaluation could include thresholds across a continuous range, along with additional performance metrics, such as F1-score or ROC curves, to provide a more nuanced understanding of the model’s accuracy at various thresholds.

In conclusion, while the authors provide an insightful prediction of PFAS occurrence in groundwater, the limited threshold analysis raises important questions about the robustness of their model’s performance. Further investigation into a broader set of threshold values is recommended to ensure the model’s applicability across different real-world scenarios. Only through such a comprehensive analysis can the model be fully validated, ensuring its reliability for use in public health and environmental monitoring.

[1] Andrea K. Tokranov et al. ,Predictions of groundwater PFAS occurrence at drinking water supply depths in the United States.Science386,748-755(2024).